DOCUMENT RESUME

ED 051 309 24 TM 000 695

AUTHOR	Webster, Murray A., Jr.; And Others
TITLE	Combining Sources of Evaluations: Six Alternative Models.
INSTITUTION	Johns Hopkins Univ., Baltimore, Md. Center for the Study of Social Organization of Schools.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	R-99
BUREAU NO	BR-6-1610
PUB DATE	Apr 71
GRANT	OEG-2-7-061610-0207
NOTE	44p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Ability, *Cognitive Processes, Comparative Analysis, Decision Making, Expectation, *Information Processing, *Information Theory, Mathematical Models, Predictor Variables, Rating Scales, *Self Concept, *Self Evaluation, Student Evaluation
IDENTIFIERS	Significant Other

ABSTRACT

Six alternative models of the ways in which individuals accept and organize information from potential sources are proposed. Parameter estimates were obtained from experiments testing the effect of evaluations of performance upon an individual's conception of his own, and others', ability to perform specific tasks. Then the models were tested against data from an independent set of experiments. Results of the tests favor a simple additive model. Possible application of the models to extensions of this experimental situation and to natural settings are discussed. Two results with implications for education are: (1) it is the perceived ability of the evaluator which determines his effectiveness and (2) evaluators with perceived low abilities tend to be more effective at lowering self-evaluation than at raising it. (DG)

ED051309

Center for Social Organization of Schools

APRIL, 1971

REPORT NO. 99

COMBINING SOURCES OF EVALUATIONS: SIX ALTERNATIVE MODELS

BY

MURRAY A. WEBSTER, JR.
S. LYNNE ROBERTS
BARBARA I. SOBIESZEK

The Johns Hopkins University

TM 000 695

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

1

STAFF

John L. Holland, Director

James M. McPartland, Assistant Director

Virginia Bailey

Thelma Baldwin

Zahava D. Blum

Judith P. Clark

James S. Coleman

David DeVries

Keith Edwards

Doris R. Entwisle

Gail Fennessey

James Fennessey

Jean Finnegan

Ellen Greenberger

Rubie Harris

Edward J. Harsch

Robert T. Hogan

John H. Hollifield

Nancy L. Karweit

Judity Kennedy

Steven Kidder

Hao-Mei Kuo

Samuel Livingston

Edward L. McDill

Rebecca J. Muraro

Jeanne O'Connor

Suzanne K. Pieper

Peter H. Rossi

Leslie Schnuelle

Aage B. Sørensen

Julian C. Stanley

Mary C. Viernstein

Diana F. Ward

Murray A. Webster

Phyllis K. Wilson

ED051309

COMBINING SOURCES OF EVALUATIONS: SIX ALTERNATIVE MODELS

Grant No. OEG-2-7-061610-0207

Project No. 61610-06-02

Murray A. Webster, Jr.
S. Lynne Roberts
Barbara I. Sobieszek

Report No. 99

April, 1971

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred.

The Johns Hopkins University

Baltimore, Maryland

INTRODUCTORY STATEMENT

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through five programs to achieve its objectives. The Academic Games program has developed simulation games for use in the classroom, and is studying the processes through which games teach and evaluating the effects of games on student learning. The Social Accounts program is examining how a student's education affects his actual occupational attainment, and how education results in different vocational outcomes for blacks and whites. The Talents and Competencies program is studying the effects of educational experience on a wide range of human talents, competencies and personal dispositions, in order to formulate -- and research -- important educational goals other than traditional academic achievement. The School Organization program is currently concerned with the effects of student participation in social and educational decision making, the structure of competition and cooperation, formal reward systems, ability-grouping in schools, effects of school quality, and applications of expectation theory in the schools. The Careers and Curricula program bases its work upon a theory of careers. It has developed a self-administered vocational guidance device to promote vocational development and to foster satisfying curricular decisions for high school, college, and adult populations.

This report, prepared as part of the School Organization program, discusses how individuals accept and organize information, and serves as a reference for social scientists to examine the utilization of expectation theory in raising the self-evaluations of school children.

ABSTRACT

Six alternative models of the ways in which individuals accept and organize information from potential sources are proposed. Parameter estimates are obtained from a set of experiments (n = 110), and the models are tested against data from an independent set of experiments (n = 114). Results of the tests favor a simple additive model. Two of the models proposed are elaborations of ideas recently proposed by Berger and Fisek (1970), and results of our tests are consistent with theirs. Possible application of the models to extensions of this experimental situation and to natural settings are discussed.

TABLE OF CONTENTS

Introductory Statement..... ii

Abstract..... iii

Self-Concept and Information Theory..... 1

The Models and the Testing..... 11

Single Source Model..... 12
Single Source Given Disagreement Model..... 15
Simple Additive Model..... 16
Averaging Effects Model..... 17
Averaging Given Disagreement, Otherwise Additive Model..... 19
Simple Operator Model..... 21

Discussion and Implications for Future Work..... 24

Implications for Education..... 33

Self-Concept and Information Theory

Self-concept, according to the tradition growing from the writings of Cooley (1902) and Mead (1934), is dependent upon the attitudes and opinions of others. If we restrict attention to evaluative aspects of the self, then this tradition would assert that the individual's high or low self-evaluation is a function of the positive and negative evaluations he receives from others. Moreover, the self-evaluation is dependent upon the evaluations received from a particular kind of others, "significant others," in Sullivan's (1947) terminology. In this paper we propose and test alternative models intended to describe more precisely the processes by which the self incorporates others' evaluations.

The proposed models are tested by application to a series of laboratory experiments in which 224 subjects participated. The experimental designs were developed in order to test propositions related to the development of conceptions of ability of self and others. One of the early experiments was a test of the proposition that a crucial determinant of whether an other could be a significant other -- that is, one whose evaluations "mattered" to the individual -- was the individual's perception of the other's ability at the task. An evaluator who was perceived to have high ability to perform the task himself was predicted to be accepted as a significant other: the individual would accept his evaluations of performance and form ability conceptions based upon them. On the other hand, if the evaluator were perceived to have low ability, the individual was predicted to be far less likely to "cognize" the evaluations, or to incorporate them into his ability conception. These predictions were confirmed in the first series of

experimental tests (Webster, 1969); evaluations from a 'high ability evaluator' were shown to affect the ability conceptions and future behaviors of subjects, while evaluations from a 'low ability evaluator' were shown to be far less likely to do so.

However, one interesting result of these first experiments was that the opinions of a low ability evaluator were not totally ignored. They were shown to be less likely to affect the cognitions and future actions of the individual, but they did produce a measurable effect in some cases. This result suggests the desirability of developing a more complete conceptual understanding of the ways in which individuals incorporate information from their environments in general, and evaluations of ability from others in particular. The perspective we adopt is one shared by a family of theories which are loosely grouped together as 'information theory.'

From an information theoretic approach, the individual is viewed as an information processing system; he is considered to perceive, interpret, and assimilate data from the external environment, and the cognitions and the future actions of the individual are assumed to be directly affected by the nature of the 'information' data which he perceives. At the same time, it is clear that not all of the "input" sensory data can be called useful "information;" in any environment, there simply are too many individual pieces of data, too many sensory stimuli, for an individual to attend to all of them, or to make sense of all of them. Therefore, he must select those data which are useful and ignore or discard the rest.

The mechanisms of this selective process are important to an understanding of cognitions and actions from an information-theoretic point of view. To understand how individuals accept and utilize information, we need answers to the following questions: Which types of information are perceived to be useful, and which are likely to be ignored? What effect does the individual source of the information have in determining the effect of the message? How does the raw "sensory input" become translated into useful "information"? What effect upon cognitions and future behavior does information have, once it has been "accepted" by the individual?

We will attempt partial answers to these general questions by addressing two more specific questions: (1) How does the individual utilize conflicting information from various sources? and (2) How does the individual utilize congruent information from two differentiated sources, or from two equivalent sources? We will examine these questions by using data from a series of similar experiments, conducted at different times, but with comparable subject pools and nearly identical experimental procedures. We will examine and test six possible models of how individuals process information.

The mechanisms by which individuals combine information have been studied by several other researchers. In work most closely related to our own, Berger and Fisek (1970) analyzed the results of experiments in which subjects were told by the experimenter that they possessed either high ability at two tasks, low ability at both, or high at one task and low at the other. Their analysis was concerned with distinguishing between what they called a 'combining' mechanism and a

'balancing' mechanism. Their results indicate that subjects seemed to utilize both possible sources of information (combining) in forming their cognitions, rather than selecting only one and ignoring the other (balancing).

Our experimental situation differs from that of Berger and Fisek in that our evaluations are not made by an "omnipotent" experimenter, so the subjects may not perceive them to be completely reliable. Thus, in addition to providing a test of similar ideas in a different setting, our experiments provide a situation in which it is possible that subjects can ignore information.

Our series of experiments were conducted as part of a program which extended and tested a theory of the effects of others' evaluations of performance upon the individual's conceptions of his own, and others', abilities to perform specific tasks. The theory and the purpose of the experiments have been reported in other works (Sobieszek, 1971; Webster, 1969, 1970; Webster and Sobieszek, 1970 and forthcoming); only those aspects relevant to the models and the model testing we will perform here will be described in this work.

The experiment consists of two phases. In Phase I, pairs of subjects enter the laboratory and perform a task which requires them to make 20 binary judgments about a series of slides. Each judgment is communicated to a third individual who has been described as possessing either unusually high ability to make these judgments, or unusually low ability. This third individual, the evaluator, supposedly decides whether each subject's judgment is correct, and communicates this information to him. Actually, the situation is controlled in a number

of ways (see Webster, 1970, for details).

At the end of the series of 20 slides in Phase I, the evaluator announces his "opinion" of the subjects' performances. He gives each subject either a high proportion of positive evaluations or a high proportion of negative evaluations. At this point, <u>if a subject believes the evaluator</u>, he should think that he is either very good at the task, or that he is very poor at the task. Thus there are two possible values for the evaluator's ability (high or low), and two possible types of evaluations given to a subject (highly positive or highly negative). The evaluator may have either high ability or low ability (the H and L conditions), and he may have told the subject that he is much better than his partner (the [+] condition) or that he is much worse than his partner (the [-] condition).

The data from the experiment are gathered in Phase II, when a second series of slides is presented. For each slide, each subject makes an 'initial choice' which is communicated to his partner. The subject then restudies the slide, and then makes a private 'final choice.' Communication is again controlled, so that subjects are told that their initial choices are in virtually continual disagreement. The basic prediction of the theory is that the higher the subject's conception of his own ability relative to that of his partner, the less likely he is to accept influence when he is told that their initial choices conflict. The relative conception of ability is predicted to be determined both by the evaluations received in Phase I, <u>and</u> by the perceived competence of the evaluator, which was announced by the experimenter. The proportion of times that subjects in each condition

resolve disagreements in favor of self, P(s), is recorded as the main statistic used to test these predictions. For example, a subject who refused to change any of his initial choices would have a P(s) = 1.00; a subject who changed 16 out of the 20 disagreeing initial choices would have a P(s) of .20.

Combining the notation for the evaluator's ability and for the nature of his evaluation for this experiment, the theory predicts the following order of experimental conditions by the P(s) statistic: H[+]>L[+]>L[-]>H[-].[1] Results from 80 subjects in this experiment -- which we will refer to as the "single-source" experiment -- are given in Table 1.[2]

TABLE 1

P(s) Values for the Single Source Experiment
(first 5 trials omitted)

Condition	n	P(s)
1. H[+]	19	.79
2. H[+]	19	.64
3. L[-]	20	.56
4. H[-]	18	.46

In terms of testing the theory which guided this research, it is significant in Table 1 that the observed ordering of conditions by P(s) data is the same as the predicted ordering.

In terms of our examination of how individuals accept and utilize information, three other facts are significant. First, the observed ordering of conditions indicates that the individual has 'accepted' and

incorporated the evaluation information in Phase I of the experiment, and has used it to determine a feature of his behavior in Phase II. Second, not only is the relative evaluation information utilized (the P(s) for respective [+] conditions is greater than the P(s) for [-] conditions), the source of the information is also utilized (the effect of either positive or negative evaluations from the H is greater than the effect of the same evaluations from the L). Third, in those conditions of the experiment where the only information available was from an individual described as possessing unusually low ability (the L conditions), subjects did not totally disregard the evaluations; the P(s) for the L[+] condition was greater than the P(s) for the L[-] condition.

Thus we identify two determinants of the effect of an evaluation upon the subject's cognitive state and upon his subsequent actions: (1) the positive or negative nature of the evaluation, and (2) the high or low ability of the person making the evaluation. In order to use this information to describe precisely how subjects utilize information from two evaluators -- both when that information is consistent, and when the two evaluators disagree with each other -- we need some estimate of the "strength of effect" of each of these factors. We need some sort of "baseline estimator" from which we can assess the additional positive or negative effects of the evaluator and of the evaluation.

Such an estimator would be the P(s) value which is derived from an experiment in which the Phase I evaluations were never communicated to subjects. They were told that their performances were being evaluated in Phase I, but that they would not be told the evaluations

until after the entire study was completed. All other details of Phase I and of Phase II were identical to the experiment previously described. This experiment may be called the 'no source' experiment, or O-S. Since the subjects do not know how they've been evaluated in the first phase of this experiment, they enter the Phase II disagreements either with unformed ability conceptions, or with conceptions determined by factors external to the experiment. The P(s) -- the proportion of times that subjects would not change their initial choices -- observed in the O-S situation is:

Condition	n	P(s)
O-S	30	.62

If we take the .62 probability from the O-S as the estimate of the baseline effect of the disagreements without any evaluations and subtract it algebraically from each condition of the single source experiment, we get a table of "evaluation weights," shown in Table 2. These weights are an estimate of the combined effect of the evaluator and the evaluation in each of our H[+], H[-], L[+], and L[-] conditions.

TABLE 2

Estimated Evaluation Weights For The Single Source Experiment

Evaluation	Weight
H+	+ .17
H-	- .16
L+	+ .02
L-	- .06

The P(s) figures observed for the first experiment and shown in Table 1 could be determined from knowledge of the 'baseline estimator' figure of .62 and the 'evaluation weights' for this experimental situation shown in Table 2. Thus, using the baseline estimator and our evaluation weights, we can predict the P(s) values from other experiments in which there are two Phase I evaluators.[3] We can predict the P(s) values in situations where these two evaluators possess either equal ability or different ability, and where they either agree or disagree as to the nature of their evaluations of performance.

It is interesting to note one feature of the estimates obtained from the single source and the O-S experiments. From the figures shown in Table 2, the effect of either a positive or a negative evaluation from a high ability evaluator is about the same: the positive evaluation from him adds .17 to the baseline estimator value, and the negative evaluation subtracts .16. However, the effect of a negative evaluation from a low ability evaluator in this situation appears to be three times as great as the effect of a positive evaluation; the negative evaluation subtracts .06 from the baseline estimator, and the positive evaluation adds only .02. It should be noted, however, that the parameters for effect of various evaluations have been estimated from minimal information; only one condition was used to obtain each estimate. Consequently it does not seem wise to attach a large amount of confidence to the exact values.[4]

Successful prediction of the results of experiments where evaluators are of equal or different ability, and where they either agree

or disagree on evaluations, depends upon the way in which individuals are believed to combine information in this situation. This combining 'process' may be represented by a variety of models, derived from a corresponding variety of substantively different theories of how information is utilized. We will consider six such models which seem to represent intuitively plausible assumptions as to how information is combined, and which make distinguishable predictions for behavior in this experimental situation.

Two additional experiments were conducted for which our models may make predictions. Both involved subjects in Phase I receiving evaluations from two evaluators. In the first experiment (n = 84), one evaluator in every group was described as possessing unusually high ability, and the other, as possessing unusually low ability. These two evaluators either agreed (conditions 1 and 5 in Table 3) or disagreed (conditions 2 and 4) on their evaluations of each subject's performances. In the second experiment (n = 30), shown as condition 3, both evaluators were described as possessing unusually high task ability, and they disagreed on their evaluations of each subject's ability. The observed P(s) values from these five conditions are shown in Table 3.

TABLE 3

P(s) Values for the Two Source Experiments

Condition	n	P(s)
1. H[+]L[+]	21	.80
2. H[+]L[-]	20	.75
3. H[+]H[-]	27	.67
4. H[-]L[+]	20	.57
5. H[-]L[-]	21	.42

The Models and the Testing

Our predictions of the P(s) values for these five conditions will depend upon the model employed to describe how individuals combine information in this situation. We will examine six models which fall into two general categories: additive models and operator models. Additive models assume that the combining process may be adequately represented by a summation of available information -- the individual "adds" units of information, which may or may not be equal in importance, to arrive at the total information which forms the basis for his action. Operator models assume that the individual performs some more complicated 'operation' upon the unit of information before assimilating it. In this experimental situation, for example, one simple additive model would assume that the individual adds up the evaluations from two sources to come to a total "weight" which he assigns to the information. A simple operator model would assume that there is some 'baseline effect', such as the P(s) from the O-S experiment, which is then proportionally altered by the effect of

evaluations and the ability of the evaluator. Therefore, the estimates derived from an operator model are a function of the baseline probability, while the estimates derived from additive models are independent of the baseline probability. Since additive models assume the more simple combining process, we will examine them first.

The general form of the additive models is shown by

$$P_o + E_{H+} + E_{L+} = P_{H+L+}$$

Where P_o = P(s) for the O-S experiment, E_{H+} = the estimated evaluation weight of a positive evaluation from the high ability evaluator, E_{L+} = the estimated evaluation weight of a positive evaluation from the low ability evaluator, and P_{H+L+} = P(s) for the H[+]L[+] condition of the two-source experiment.

When the individual is aware that more than one other individual is giving him opinions, a first step in utilizing this information is to decide whether to accept all available data, or to "perceive it selectively" and ignore some of it. In this experimental situation, the question is whether a subject who is receiving evaluations from two evaluators will "pay attention" to both of them, or whether he will simplify the situation cognitively by choosing only one of them. The first model we will examine tests the latter process.

1. The Single Source Model

The single source model assumes that in the case of more than one potential source of evaluation the individual will accept information from only one. Thus it becomes necessary to specify which one will be accepted and which ignored.

A basis for making this specification is provided in the original theory which guided these experiments, and which was supported in the results of the original single source experiments shown in Table 1. The high ability evaluator was shown to be more effective than the low ability evaluator. Therefore, if our model asserts that the individual will accept only one potential source of information, when both a high ability evaluator and a low ability evaluator are available, it is reasonable to expect that the individual will accept the high ability evaluator and will ignore the low ability evaluator. For these experiments, this process would have two consequences: First, the P(s) from the H[+] condition would be the same as the P(s) from the H[+]L[+] and the H[+]L[-] conditions. Second, the P(s) from the H[-] condition would be the same as the P(s) from the H[-]L[-] and the H[-]L[+] conditions. In each case, the L[+] or L[-] would be ignored.

For the case of two disagreeing high ability evaluators -- the H[+]H[-] experiment -- it is difficult to predict which of them would be accepted. To obtain a simple estimate of P(s), we can assume that half the subjects in this experiment will accept only the positive evaluations of self and half will accept only the negative evaluations of self. Those who accept the H[+] evaluation will display the same P(s) as subjects who received positive evaluations of self from only one high ability evaluator in the single source experiments. Those who accept the H[-] evaluation will display the P(s) of subjects in the comparable condition in the single source experiment. The respective P(s) figures for these two conditions

are .79 and .46, which we can average to obtain an estimate of .63 for the H[+]H[-] experiment from this model.

The P(s) figures predicted by this single source model are shown in Table 4, with the figure for the H[+]H[-] experiments in brackets, to indicate that an additional assumption was required to get the predicted figure.

TABLE 4

The Single Source Model:
Expected and Observed P(s) Values

Condition	Expected P(s)	Observed P(s)	Discrepancy (|exp.-obs.|)
1. H[+]L[+]	.79	.80	.01
2. H[+]L[-]	.79	.75	.04
3. H[+]H[-]	[.63]	.67	[.04]
4. H[-]L[+]	.46	.57	.11
5. H[-]L[-]	.46	.42	.04

The general additive model

$$P_o + E_{H+} + E_{L+} = P_{H+L+}$$

was used for calculating the predicted figures, with the restriction for the two evaluation system that $E_{L+} = E_{L-} = 0$. Therefore,

$$P_o + E_{H+} = P_{H+L+}.$$

Comparison of predicted with observed figures in Table 4 indicates that this first model provides a close prediction of the P(s) for condition 1, and fairly close predictions for conditions 2, 3, and 5. The prediction is less close for condition 4.

2. The Single Source Given Disagreement Model

This model, a variant of the first, assumes that when all sources are giving consistent information, the tendency of the individual is to accept more than one source. When potential sources disagree in their information, then the individual will choose one of them and disregard the other, as in the single source model. More precisely, this second model assumes that in conditions 1 and 5 of the experiments, the individual will follow an additive process for consistent information; he will add the bits of information from each potential source. In conditions 2, 3 and 4 -- where sources disagree -- the individual will react as he did for the single source model and accept only one source. Table 5 presents the predicted and observed P(s) values for this second model.

TABLE 5

The Single Source Given Disagreement Model:
Expected and Observed P(s) Values

Condition	Expected P(s)	Observed P(s)	Discrepancy (|exp.-obs.|)
1. H[+]L[+]	.81	.80	.01
2. H[+]L[-]	.79	.75	.04
3. H[+]H[-]	[.63]	.67	[.04]
4. H[-]L[+]	.46	.57	.11
5. H[-]L[-]	.40	.42	.02

This elaborated model provides a good fit to the observed data for condition 1, and a better fit than the single source model for condition 5, while retaining the same predictions of conditions 2, 3

and 4 as the single source model. The single source given disagreement model, then, provides a better overall fit than the simple single source model, the simplest of the 'balancing' models.

3. The Simple Additive Model

This model assumes that the individual accepts all information available in the situation and simply sums it in order to arrive at a useful conclusion. The individual adds up all available information, both when sources agree and when they disagree, in the manner postulated only for cases of agreeing sources in model #2. The formula for calculating predicted values for the simple additive model is the same as the general additive model, with no restrictions: $P_o + E_{H+} + E_{L+} = P_{H+L+}$. Table 6 reports the predicted and observed P(s) values, using the simple additive model.

TABLE 6

The Simple Additive Model:
Expected and Observed P(s) Values

Condition	Expected P(s)	Observed P(s)	Discrepancy (| exp.-obs. |)
1. H[+]L[+]	.81	.80	.01
2. H[+]L[-]	.73	.75	.02
3. H[+]H[-]	.63	.67	.04
4. H[-]L[+]	.48	.57	.09
5. H[-]L[-]	.40	.42	.02

The simple additive model gives a better overall fit to the observed data than either the single source model or the single

source given disagreement model. It retains model #2's good predictions of conditions 1 and 5, provides the same fit as models #1 and #2 for condition 3, and has a better fit for conditions 2 and 4 than either previous model. Also for condition number 3, the predicted P(s) and observed discrepancy are no longer bracketed, since both the H[+] and the H[-] are weighed in this model.

In addition, we may note the greater conceptual simplicity of the simple additive model compared to the single source given disagreement model. Both on grounds of theoretical simplicity and on grounds of accuracy of empirical prediction, the simple additive model is to be favored over either of the two 'balancing' models.

4. The Averaging Effects Model

Our fourth model asserts that the total effect of each additional potential piece of information is diminished by the effect of every previous piece of information -- in other words, that information from more than one of several equivalent sources is 'averaged' to produce the final cognition. In our experimental situation, this model would predict that the effect of a negative evaluation from a low ability evaluator would decrease the strength of effect of a positive evaluation from a high ability evaluator. The general form of an averaging model is:

$$P_{H+L+} = P_o + \frac{E_{H+} + E_{L+}}{k},$$

where k is the number of potential sources of information, and the other terms are interpreted as in the previous models.

In any empirical operation, the averaging model seems most plausible for situations where the available information is inconsistent. But it does not seem likely that consistent information would be averaged. To assert that individuals average information from one high ability evaluator and one low ability evaluator when the evaluators agree is to assert that the agreement from the low ability evaluator in some way decreases the subjects' confidence in the information from the high ability evaluator.

Although it seems intuitively unlikely, this assertion can be plausibly argued. The low ability evaluator in these experiments is expected to make some mistakes in his evaluations, and if the high ability evaluator always agrees with him, then the high ability evaluator may be making a few mistakes also. Thus, the agreement of a low ability evaluator may be a factor in reducing the credibility of the high ability evaluator.

We first make the assumption, then, that the averaging model may be applied to all cases. Table 7 presents a comparison of the observed P(s) values and those which are predicted from a simple averaging model.

TABLE 7

The Simple Averaging Model:
Expected and Observed P(s) Values

Condition	Expected P(s)	Observed P(s)	Discrepancy (| exp.-obs.|)
1. H[+]L[+]	.72	.80	.08
2. H[+]L[-]	.68	.75	.07
3. H[+]H[-]	.63	.67	.04
4. H[-]L[+]	.55	.57	.02
5. H[-]L[-]	.51	.42	.09

Compared to the simple additive model, the simple averaging model gives the same prediction for condition 3 and a better prediction for condition 4. However, the simple additive model gives better predictions for conditions 1, 2, and 5. Thus the simple averaging model is better than the simple additive model for only one case; however, that case is the one for which the simple additive model gives a particularly poor prediction.

5. The 'Averaging Given Disagreement, Otherwise Additive' Model

In the previous averaging model, we assumed the condition that averaging occurs in all cases. For model #5 we assume the more intuitively likely condition -- that the individual averages information only when he is exposed to inconsistency; that is, in those cases of the experiment where the evaluators disagree. We make the further assumption that the simple additive model should be applied in cases of agreement, since it provided that best fit to those conditions. Table 8 presents the results of predicted and observed

P(s), using a model which assumes a simple additive process when information is consistent, and averaging when it is inconsistent.

TABLE 8

The Averaging Given Disagreement Otherwise Additive Model: Expected and Observed P(s) Values

Condition	Expected P(s)	Observed P(s)	Discrepancy (|exp.-obs.|)
1. H[+]L[+]	.81	.80	.01
2. H[+]L[-]	.68	.75	.07
3. H[+]H[-]	.63	.67	.04
4. H[-]L[+]	.55	.57	.02
5. H[-]L[-]	.40	.42	.02

The 'averaging-given-disagreement-otherwise-additive' model retains the good prediction of condition 4 from the simple averaging model, and it incorporates the good predictions of conditions 1 and 5 from the simple additive model. All predictions from this model are reasonably close to the observed data. Compared to the predictions from the simple additive model, however, predictions from the averaging-given-disagreement are not significantly better. The simple additive model makes a considerably better prediction for condition 2, while the averaging-given-disagreement makes a better prediction of condition 4. Both models make the same prediction for condition 3. Therefore, on the criterion of empirical prediction, the choice between these models is not clear-cut.

For simplicity, the simple additive model is clearly preferable. The averaging-given-disagreement model requires that an additional

assumption be made, and moreover, it provides no theoretical reason for using the simple additive model in cases where the available information units are consistent. For the tests shown in Table 8, the simple additive model was used for the consistent evaluation conditions on the ad hoc grounds that it provided a good fit to observed data, but acceptance of the 'averaging-given-disagreement-otherwise-additive' model would require a more complete justification for doing this. Such a justification, of course, remains to be worked out. Therefore, on grounds of theoretical simplicity, and expecially in view of the fact that the simple additive model provides a reasonably good fit of the data, we conclude that it is to be preferred over the averaging-given-disagreement model.

6. A Simple Operator Model

An operator model assumes that an individual processes information through a multiplicative rather than an additive process. This model takes the O-S probability of rejecting influence (.62) as a baseline. This is multiplied by an estimate of the multiplicative effect of positive or negative evaluations from a high ability or a low ability evaluator to obtain the observed P(s) of the condition. These estimates of multiplicative effect will be represented by alphas.

In order to obtain the estimate of alpha for a positive evaluation from a high ability evaluator, for example, we take the O-S P(s) figure as the estimate of baseline effect of no evaluator and then determine the fiugre by which it must be multiplied in order to obtain the observed P(s) for the H[+] condition.

Thus, the formula for obtaining the alpha-estimate for the H+ condition is

$$\alpha_{H+} \cdot P_o = P_{H+} \text{ or } \alpha_{H+} = \frac{P_{H+}}{P_o}$$

Using this general formula and solving for the alpha estimates of the H[+], L[+], L[-] and H[-] conditions, we get the values shown in Table 9.

TABLE 9

Alpha-Estimates for the Operator Model

Parameter	Estimate
α_{H+}	1.274
α_{H-}	.742
α_{L+}	1.032
α_{L-}	.903

These alpha estimates are dependent upon the O-S probability. They are estimates of the proportional increase in the initial O-S probability which is caused by the high or low ability evaluator and by the positive or negative evaluation. In comparison, the additive models use estimates of effect of evaluation which are independent of the O-S probability. That is, one may speak of the effect of a positive evaluation from a high ability evaluator as having a value of +.17 (shown in Table 2), and this value could be added to whatever the baseline probability for no evaluator might be in a particular experimental situation.

This is not the case for the operator model. The alpha-estimates, since they speak of a proportional increase over the given baseline probability for the given situation, are specific to the experimental situation and to the subject pool used. One could not use alpha-estimates from one experimental situation in a different situation, nor use alpha-estimates from one subject pool in another subject pool.

The general form for the operator model is given by:

$$\alpha H+ \cdot \alpha L+ \cdot P_o = P_{H+L+},$$

where P_o is the P(s) for the O-S experiment, and P_{H+L+} is the predicted value for the H[+]L[+] experiment. Using this formula, the P(s) values shown in Table 10 would be predicted for the two evaluator experiments.

TABLE 10

The Operator Model:
Expected and Observed P(s) Values

Condition	Expected P(s)	Observed P(s)	Discrepancy (| exp.-obs. |)
1. H[+]H[+]	.82	.80	.02
2. H[+]L[-]	.71	.75	.04
3. H[+]H[-]	.58	.67	.09
4. H[-]H[+]	.47	.57	.10
5. H[-]L[-]	.41	.42	.01

The operator model does not do a particularly satisfactory job of predicting the observed values for the two evaluator experiments. Only for conditions 1 and 5 are the predictions reasonably close to the observed figures; for conditions 3 and 4

the fit is poor, and for condition 2 it is moderate. Compared to the additive models, the operator model does less well than any of them at predicting the observed figures. In particular the fit by the operator model is considerably poorer than it was by the simple additive model, the one which generally made the best predictions. Based upon these comparisons, we conclude that the additive models as a set, and especially the simple additive model, are preferable to the operator model both for conceptual simplicity and adequacy of empirical predictions.

Discussion and Implications for Future Work

The results of our model testing support the following conclusions. First, among the additive models, the simple additive model is preferable. It is the simplest model and the one which makes the most satisfactory predictions of observed data. Second, the additive models as a group are preferable to the operator model for the same reason: the additive models are simpler, and they enable better predictions of the observed data.

In terms of the more general question of how individuals accept information from a variety of potential sources, many theories of cognitive consistency argue that individuals will accept consistent information, but will distort information which is inconsistent (Heider, 1946; Secord and Backman, 1961). On the other hand, other information processing theories assert that the individual accepts all types of information without regard to consistency, and that his final set of cognitions is partially determined by the total of available information. In our studies, the former process is

represented by the 'single source' and the 'single source given disagreement' of the additive models; the latter, by the 'simple additive' and the 'averaging' models. The results of the comparisons of predicted and observed effects support the latter models and thus, the latter process.

Recent experimental work by Berger and Fisek (1970) also has attempted to differentiate between these two processes (which they called, respectively, 'balancing' and 'combining'). Their results were partially consistent with ours; they found support for the combining process. However, Berger and Fisek reported results which were restricted to cases in which individuals had conflicting information from two sources who were identical in 'plausibility,' or in likelihood of being accepted. In our experiments, the 'plausibility' of the low ability evaluator may be considered to be lower than that of the high ability evaluator. The advantage of this differential is that it is possible to distinguish the results of 'averaging' in our experiments from the results of ignoring both potential sources of information, something which Berger and Fisek note was not possible in their experiments reported to date. For example, the P(s) for the H[+]L[+] condition shown in Table 3 differs from both the P(s) for the H[+] condition and the H[+]L[-] condition. This is consistent only with the interpretation that subjects accept information from both potential sources and then combine it.

Our results also differ slightly from those of Berger and Fisek in that they tested only one model of each type (combining and

balancing), while we tested several, slightly differing versions. In general, our results are consistent with theirs: the evidence for some sort of combining process was stronger than the evidence for a balancing process. The model which Berger and Fisek proposed and for which they found support seems most closely related to our 'averaging' model (#4); that is, the model which asserts that individuals will accept several sources of information and combine them by weighing each bit of information. While this model made fairly good predictions to our observed data, the simple additive model (#3) made even better predictions. The simple additive model proposes that the individual combines information, not by weighing it, but simply by summing bits of contradictory or consistent information. The processes assumed in both models are roughly similar. However, it is possible to distinguish between the exact form of the process expected from the two models, and our results provide the better support for model #3. Thus it seems accurate to consider our results as a refinement of the theoretical ideas proposed by Berger and Fisek.

In the future work, these models might be applied to a third series of experiments in which the number of evaluators, the perceived ability of the evaluators, or the nature of the positive and negative evaluations from the evaluators are systematically varied. For example, predictions may be extended from any of these models to cases of more than 2 evaluators, and these could be tested empirically by adding an evaluator to the basic experimental situation. Experimental tests of this sort would have three evaluators of subjects in Phase I of the experiments, and the same series of critical trials in Phase II for collecting data.

Two considerations would have to be taken into account in such 'three evaluator' experiments. First, the parameter values used in some of these models would have to be altered for three evaluator experiments, or some constant factor would have to be introduced in order to prevent empirically impossible predictions. Using the present evaluator effect estimates and alpha estimates, several of these models would predict P(s) values of less than 0 or of more than 1.0 for cases of three high ability evaluators. In fact, the alpha estimates used for the operator model (#6) would predict a P(s) of more than 1.0 for a two evaluator experiment in which two high ability evaluators both gave the subject positive evaluations. The evaluation effect estimates used would predict a P(s) of .96 from a simple additive model (#3) for the same H[+]H[+] experiment. Therefore, application of these models to the three evaluator situation would require some adjustment of the parameters or of the models in order to avoid such empirically meaningless outcomes.

Second, application of these models to the three evaluator situation would have to take account of a 'ceiling effect' and a 'floor effect' in the observable P(s) values for this experimental situation. In view of past work on these experiments, it seems likely that P(s) values above .80 (ceiling) or below .20 (floor) are very unlikely to be observed in this situation. One reason for the 'floor' can be appreciated in terms of 'person orientation,' (Webster, 1970), which is the desire to have the final choices in this situation reflect the subject's <u>own</u> ability, rather than a willingness to obtain the correct answer regardless of whether it

is the subject's own answer or his partner's. To a noticeable extent in these experiments, subjects display resistance on any given trial to accept an answer from their partner, even when they believe it is more likely to be correct than their own initial choice. Such resistance is sometimes explained by subjects in the post-session interviews as "I like to make my own decisions," or "I stick to my own choice, even though it may be wrong." Experimental designs may be altered in the attempt to minimize person orientation, but it has not yet been possible to eliminate its effects from the situation.

One reason for the 'ceiling' effect (as well as an additional reason for the 'floor' effect) may be understood when the operational meaning of a P(s) of .80 is considered: In an experiment with 20 critical trials, a P(s) of .80 means that the subject has changed his initial choices on only 4 slides. While occasionally it is observed that subjects will change even fewer than 4 initial choices, in any group where individuals are oriented to teamwork, or where they have a collective orientation to the task, it is unlikely that an individual will change fewer than four of this initial choices; this would seem to require an almost total disregard for the partner's opinions.

With these qualifications in mind, predictions from these models could be made to the three evaluator cases; for example to:

1. H[+]H[+]L[+]
2. H[+]H[+]L[-]
3. H[+]H[+]H[-]
4. H[+]H[-]H[-]

Comparison of cases 3 and 4 would be interesting in consideration of the 'combining' vs. 'balancing' issue in accepting information. If the balancing approach were to be supported, subjects in case 3 would be expected to form high self-evaluations which would be equivalent to the evaluations formed in the H[+] condition of the single evaluator experiment in the first series; and subjects in case 4 would be expected to form low self-evaluations which were equivalent to those formed in the H[-] condition of the single source experiments. If the combining approach were to be supported, subjects in case 3 should form lower self-evaluations than those in the equivalent conditions of the single evaluator experiments, and those in case 4 should form higher self-evaluations than those in the equivalent condition of the single evaluator experiment.

Comparisons could also be drawn between cases 2 and 3 above, for the same purpose. If the 'single source' formulation were to be supported, the P(s) for cases 2 and 3 should be the same. If an averaging or a simple additive model were to be supported, they should differ with case 2 > case 3.

The intuitive appeal of the single source model described in this paper rests upon the belief that an individual confronted with a variety of potential sources of information must choose between them on some basis. In the case of a large number of possible sources of information, it seems intuitively plausible to expect that the individual will simply disregard some sources entirely. Our test of the single source models for the two evaluator case indicates that individuals in this situation can attend equally to two evaluators

at once. Establishing the point at which selectivity becomes noticeable would be important for extending either an additive model or an operator model of combining information. This would be especially true if the threshold were low, as it would be, for example, if some disregarding became noticeable with only three evaluators.

Finally, the numerical values used in this work for parameter estimates of effect of evaluation for the additive models merit some discussion. First, comparisons may be drawn between the estimates of effect of a positive evaluation vs. a negative evaluation from the same evaluator, by comparing the estimates of the H+ to H- and the estimates of L+ to L- shown in Table 2. The estimates of effect of evaluations from a high ability evaluator do not differ appreciably: positive evaluation is predicted to raise the P(s) by .17; and negative evaluation, to lower the P(s) by .16. However the estimates of effect of evaluations from the low ability evaluator do differ appreciably: positive evaluations are predicted to raise the P(s) by only .02, and negative evaluations, to lower the P(s) by .06. In absolute terms, the difference between .02 and .06 is small, and it is possible to ascribe it to chance factors operative in this experimental sitaution. But these estimates enable quite good predictions for the second series of experiments, in which there were evaluations from a low ability evaluator. Moreover, though the absolute difference between .02 and .06 is small, the proportional difference is considerable. Therefore, we may ascribe some reliability to these figures, and may consider an implication of this proportional difference.

The idea is frequently expressed in social psychology that individuals in some sense 'prefer' a positive self-image, and thus that they will act, either behaviorally or cognitively, to maximize the actual or the perceived positive evaluations they receive from others. One way of maximizing positive evaluations received from others is to emit behaviors valued by those others; for example, Zetterberg (1957) has postulated that individuals will give 'compliance' in return for positive evaluations. A second way to maximize positive evaluations is cognitively to distort evaluations from others; either to change the valence of negative evaluations, or to ignore a greater proportion of negative than of positive evaluations. The behavioral means of seeking positive evaluations was not available to subjects in our experiments, but the cognitive means was available.

For the high ability evaluator cases, it may not be surprising that the effects of the positive evlautions and the negative evaluations were virtually identical. The high ability evaluator is one whose opinions would be difficult to ignore from an theoretical or intuitive viewpoint. But there is some reason to believe that the low ability evaluator may be overlooked by some subjects. Indeed, this is a central assertion of the theory under test in these experiments: the low ability evalutor is perceived to be less competent to evaluate performance, and thus he is less likely to become a 'significant other' whose evaluations become part of the self-evaluation of the individual.

In view of these considerations, it is striking that it is the positive evaluations from the low ability evaluator which have the smaller effect in this situation. Whether his positive evaluations

are ignored by a greater proportion of subjects, or whether all subjects ignore a greater proportion of his positive evaluations (an independent question which cannot be answered from the data we have presented here), it is the positive evaluations, not the negative evaluations, which are the more ignored. If confidence may be placed in our estimated values, positive evaluations are ignored by a ratio of 3 to 1 over negative evaluations. This interpretation, of course, is directly counter to an asserted tendency to maximize the self-image.

Also, one may compare the relative 'strengths of effect' of the high and low ability evaluators. If we consider the estimates of .16 and .17 to be reasonably accurate for the high ability evaluator, and the accurate estimate for the low ability evaluator to be somewhere between .02 and .06, then it is evident that the high ability evaluator is about three to nine times as effective as the low ability evaluator in this situation. One simple interpretation of this finding is that 'experts' are considerably more influential than 'non-experts' in determining an individual's estimate of his ability.

The large difference in our estimates of effect of the high and the low ability evaluators makes the routine practive of soliciting multiple evaluations from others of doubtful competence appear unsound. In these experiments, for example, if one adopts the model which did the best job of predicting data from the two evaluator experiments, the simple additive model, then the effect of one single high ability evaluator is equal to the effect of from 3 to 9 low ability evaluators. Especially in the example of a school child, where the desire may well be to <u>raise</u> his self-evaluation, our estimates indicate that it would require uniformly positive evaluations from <u>nine</u> separate

low ability evaluators to equal the effects of a single high ability evaluator. Of course the probability that nine individuals will give uniformally positive evaluations of any individual's performance is likely to be rather low.

Implications for Education

Children in school systems are regularly confronted with evaluations from a wide variety of sources of their performance in different areas. Their academic performance is evaluated by a homeroom teacher and by teachers in special subjects, as well as teaching aides or student teachers. Substitute teachers occasionally take over classes and are called upon to evaluate a child's written or verbal performance. In non-academic areas such as athletics or social skills, children are likely to receive evaluations not only from those who evaluate as part of their role in the formal school system (the teaching personnel), but also from other students, and from adults such as parents. The ways in which a child makes use of opinions and information from these various sources will be major determinants of his self-evaluation at the various tasks and activities which comprise school experiences.

A result which consistently appears in all of these experiments (as well as in previous reports of Source experiments; for example, in Webster, 1970; Sobieszek, 1971; Webster and Sobieszek, 1970) is that it is the perceived ability of the evaluator, not his formal authority or evaluative position within the social system, which determines his effectiveness. If the subject feels that the evaluator has low ability, even when his evaluations are the only available

information in the situation, the evidence is that they will be largely unaccepted. Generalization of this finding may be made to the situation of the teacher who is believed by the student to have low ability at the subject he is teaching; for example, the mathematics teacher who is suddenly called upon to substitute in English and who misuses grammar. It also may be analogous to the situation of a nervous teacher who, though highly competent in her subject, conveys to the students the impression that she does not know her subject well. The evidence from our experiments is that such individuals will be ineffective in forming or in changing students' conceptions of their own abilities.

The ways in which individuals utilize multiple sources of information, both when the potential sources agree and when they disagree, are assumed to be different in the different models examined in this work. Two results of the comparisons may be applied to the school setting. First, it was found that both for cases of agreement and disagreement between evaluators, a simple additive model gave the most satisfactory predictions of observed results. This means, for example, that a student who is evaluated differently by two others (two teachers, two other students, or one teacher and one student) will accept evaluations from both of them. In cases where the sources agree on their evaluations, the effects of more than one source will be greater than the effect of one. Positive evaluations of classroom performance from two or more teachers will 'cumulate' to produce a higher self-evaluation than would be produced as the result of evaluations from only one teacher. In cases where the sources disagree, the effect also will be cumulative; the student

will not choose to accept only one of them and ignore opinions from the others. However, it should be remembered here that the perceived ability of the evaluators is extremely important in determining their relative effects.

Second, the differential effect of positive and negative evaluations from low ability evaluators may have important consequences for the classroom. The effects of either positive or negative evaluations from a high ability evaluator are about equal; a teacher perceived to be highly competent would be about equally effective at raising or lowering a child's self-evaluation. However, an evaluator of low perceived ability, such as a teacher's aide, may well have differential effectiveness at raising or lowering the child's self-evaluation. The evidence from these studies is that the low ability evaluator is roughly three times as effective at lowering self-evaluations as he is at raising them. This implies that a teacher's aide could not have much effect in raising a child's self-evaluation in a classroom subject, but might have a noticeable effect at lowering it. Put in more general terms, the results of this model testing suggest that it may be easier for individuals of low ability to lower the self-evaluations of other than to raise them. Since many of the behaviors and attitudes which are considered to contribute to learning and to academic achievement are positively correlated with a positive self-image (for examples, see Entwisle and Webster, 1970), this finding indicates a potential danger to children, and steps may need to be taken to guard against the unintended effect of evaluations from such individuals.

FOOTNOTES

[1]The explicit derivation of this prediction is contained in the earlier report (Webster, 1969). The exact theoretical reasons for the prediction are not central to this report and so are omitted; however it should be clear from the predicted ordering that P(s) is expected to vary directly with two factors: (1) the ability of the evaluator, and (2) the relative proportion of positive evaluations given to the subject.

[2]There are 20 critical (disagreement) trials in all these experiments. For all tables in this report, the P(s) were calculated for the last 15 disagreement trials only, since there is reason to believe that the first few trials of the disagreement phase are used by the subject to adjust to an unfamiliar situation. Both variance across trials and variance across subjects are higher for the first block of 5 trials in these experiments than for any subsequent block of trials. Also, there is no theoretical reason to expect a systematic change in results as a function of the disagreement resolution process, nor is there evidence of such a change in these experiments. For these reasons, we conclude that the final three blocks of trials provides the most representative and stable estimate of P(s) for each condition.

Also, data from about 5%-10% of subjects in these experiments were excluded from analysis, on grounds of considerable evidence that subjects did not meet one or more conditions of the experimental design. For example, subjects who clearly did not believe that the disagreements were real could not be said to be accepting or rejecting influence in making their final choices, and so were excluded. Complete criteria for exclusion as well as details on how the inclusion/exclusion decisions were made are contained in Webster, 1970.

[3]Because we make direct comparisons of data across these various experiments, a word about comparability of the situations and of the subject pools is in order. The designs of the various experiments are all variants of a basic experimental situation developed at Stanford by Joseph Berger and his associates. Differences between the experiments are slight, being limited to the changes required for tests of the different versions of the theory. The differences probably would not be noticeable to an observer who was not closely acquainted with the particular theory derivations under test.

Subjects for the experiments reported here were all volunteers recruited from English courses at a California Junior College. At the time of the experiments, none was under 17 years of age, and none was over 24. The single source experiments were conducted in spring, 1968, with male subjects. With the exception of the

H[+]H[-] condition, the two source experiments were conducted in spring, 1969, with female subjects. The no source and the H[+]H[-] experiments were conducted during the summer, 1970, with female subjects.

Though the time difference may have produced slight differences in the subject pools, in terms of any of the frequently used measures -- such as academic ability, SES of families, etc. -- there was no change during the time these experiments were conducted. All subjects were recruited from regular classes at the college, not from evening or summer sessions. The sex difference between the single source experiments (male subjects) and the later experiments (female subjects) conceivably could make direct comparisons of data dangerous, but in the absence of specific information regarding the effect of sex in these experiments, we have chosen to treat this difference as inconsequential.

[4]The self-evaluation theory which guided these experiments does not give any reason to expect that under some circumstances negative evaluations would produce greater effect than positive evaluations from the same evaluator, nor is there any feature of this general experimental design which is clearly related to the observed effect. Therefore it might seem reasonable to impose the added condition that the effects of positive and negative evaluations from the same evaluator have the same degree of effect, or more formally: $|H+| = |H-|$ and $|L+| = |L-|$. This restriction would offer the advantage of increasing the amount of information used to estimate the 'evaluation effect' parameters, for two conditions would be used for each. As the simplest way of using two conditions, the effects of the positive and negative evaluations from the same source could be averaged, to obtain the following values:

Evaluation	Weight
H(+ or -)	.17
L(+ or -)	.04

However, there is no readily apparent justification for adding such a restriction to parameters, and use of the values shown above would not change the results of any of the comparisons to be made later in this paper. Therefore, we have decided to use the simplest estimates of the evaluation effect parameters, those shown in Table 2, and to note that they were estimated on the basis of minimal information (usually, about 20 subjects for 15 trials, or 300 units of information).

REFERENCES

Berger, J. and M. H. Fisek. Consistent and inconsistent status characteristics and the determination of power and prestige orders. Sociometry, 1970, 33, 3, 287-304.

Cooley, C. H. Human Nature and the Social Order. New York: Charles Scribner's Sons, 1902.

Entwisle, D. R. and M. Webster. Raising children's expectations for their own performance. Baltimore: Report No. 87, Center for Social Organization of Schools, The Johns Hopkins University, 1970.

Heider, F. Attitudes and cognitive organization. Journal of Psychology, 1946, 21, 107-112.

Secord, P. F. and C. W. Backman. Personality theory and the problem of stability and change in individual behavior: an interpersonal approach. Psychological Review, 1961, 68, 21-32.

Sobieszek, B. Multiple sources and the formation of performance expectations, to appear in Pacific Sociological Review, 1971.

Sullivan, H. Conceptions of Modern Psychiatry. Washington, D.C.: W. H. White Psychiatric Foundation, 1947.

Webster, M. Source of evaluations and expectations for performance. Sociometry, 1969, 32, 3, 243-258.

Webster, M. Status characteristics and sources of expectation. Baltimore: Report No. 82, Center for Social Organization of Schools, The Johns Hopkins University, 1970.

Webster, M. and B. Sobieszek. No-source and two-source experimental study. (Mimeo.), Baltimore: Department of Social Relations, The Johns Hopkins University, 1971.

Webster, M. and B. Sobieszek. Sources of Self-Evaluation, research monograph, forthcoming.

Zetterberg, H. L. Compliant actions. Acta Sociologica, 1957, 2, 179-201.